Structural analysis of Salmonella enterica effector protein SopD

Available online at www.sciencedirect.com
SCIENCE DIRECT®
Biochimica et Biophysica Acta 1698 (2004) 219–226
BIOCHIMICA ET BIOPHYSICA ACTA
BBA
www.bba-direct.com

Structural analysis of *Salmonella enterica* effector protein SopD

Michael W. Wood[a,1], Christopher Williams[b,1], Abhishek Upadhyay[b], Andrew C. Gill[c], Didier L. Philippe[b], Edouard E. Galyov[a], Jean M.H. van den Elsen[b,*], Stefan Bagby[b,*]

[a] *Division of Environmental Microbiology, Institute for Animal Health, Compton Laboratory, Berkshire RG20 7NN, UK*
[b] *Department of Biology and Biochemistry, University of Bath, Bath BA2 7AY, UK*
[c] *Division of TSEs, Institute for Animal Health, Compton Laboratory, Berkshire RG20 7NN, UK*

Received 6 August 2003; received in revised form 28 November 2003; accepted 1 December 2003

Available online 16 December 2003

Abstract

Salmonella outer protein D (SopD) is a type III secreted virulence effector protein from *Salmonella enterica*. Full-length SopD and SopD lacking 16 amino acids at the N-terminus (SopDΔN) have been expressed as fusions with GST in *Escherichia coli*, purified with a typical yield of 20–30 mg per litre of cell culture and crystallized. Biophysical characterization has been carried out mainly on SopDΔN. Analytical size exclusion chromatography shows that SopDΔN is monomeric and probably globular in aqueous solution. The secondary structure composition, calculated from the CD spectrum, is mixed (38% α-helix and 26% β-strand). Sequence analysis indicates that SopD contains a coiled coil motif, as found in numerous other type III secretion system-associated proteins. This suggests that SopD has the potential for one or more heterotypic protein–protein interactions. Limited trypsin digestion of SopDΔN, monitored by both one-dimensional proton NMR spectroscopy and SDS-PAGE, shows that the protein has a large, protease-resistant core domain of 286 amino acid residues. This single-domain architecture suggests that SopD lacks a cognate chaperone. In crystallization trials, SopDΔN produced better crystals than either full-length SopD or trypsin-digested SopDΔN. Diffraction to 3.0 Å resolution has so far been obtained from crystals of SopDΔN.

Keywords: *Salmonella*; SopD; Type III secretion

1. Introduction

Serovars of the Gram-negative intracellular pathogen *Salmonella enterica* are a major economic and welfare problem affecting both man and food-producing animals. They cause a wide range of diseases from self-limiting intestinal infections to severe systemic diseases, such as typhoid fever [1,2]. Key to the bacterium's pathogenesis are two type III secretion systems (TTSSs) [3,4]. TTSSs resemble molecular syringes for the injection of multiple bacterial virulence effector proteins into the host cell cytoplasm that modify host cell physiology to the benefit of the pathogen. The supramolecular structure of the *Salmonella* TTSS resembles the flagellar basal body and comprises two parts: (i) a 7–8 nm wide, 60 nm long external needle; and (ii) a shorter cylinder formed by plates (20–40 nm in diameter) that presumably traverse both bacterial membranes and the peptidoglycan layer [5,6]. Similar structures have also been observed in *Shigella flexneri* [7,8]. *Salmonella* has two TTSSs: on contact with the host cell, TTSS1 injects numerous virulence effector proteins that subvert diverse signal transduction pathways resulting in bacterial uptake and enteritis, whereas TTSS2 is involved in survival of the bacterial cell once inside the host cell.

One of the effector proteins secreted by *Salmonella* TTSS1 is *Salmonella* outer protein D (SopD). SopD has been shown to act in a concerted manner with another *Salmonella* TTSS1 effector, SopB, to promote inflammatory responses and fluid secretion in *Salmonella dublin*-infected calf intestines [9]. Along with SipA, SopA, SopB and SopE2, SopD has also been implicated in induction of diarrhea during *Salmonella typhimurium* infection of calves [10]. The SopD amino acid sequence is highly conserved among different *Salmonella* serovars, particularly serovars *dublin*, *typhi* and *typhimurium* (98% sequence identity across the complete

* Corresponding authors. Tel.: +44-1225-386436; fax: +44-1225-386779.
E-mail addresses: bssjmhve@bath.ac.uk (J.M.H. van den Elsen), bsssb@bath.ac.uk (S. Bagby).
[1] These authors contributed equally to this work.

1570-9639/$ - see front matter © 2003 Elsevier B.V. All rights reserved.
doi:10.1016/j.bbapap.2003.12.003

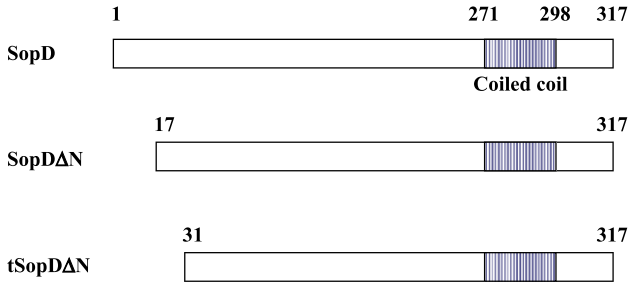

Fig. 1. Schematic representation of the SopD constructs generated in this study.

SopD amino acid sequence). Incomplete sequences are available for SopD from several other *Salmonella* serovars. The sequence identity between any of these incomplete sequences and the three complete SopD sequences (serovars *dublin*, *typhi* and *typhimurium*) is at least 69%.

No other sequence with obvious homology to the *sopD* gene product sequence has been found among the GenBank-EMBL entries, making it difficult to suggest a possible function for SopD. Analysis of the genome sequence of *S. typhimurium* [11] identified a homologue of SopD termed SopD2, but even though SopD and SopD2 share 42% sequence identity over their entire length, they seem to have different intracellular sites of action: SopD2 was found to be membrane-associated, whereas SopD localized to the cytosol [12] (the precise targeting of SopD within the cytosol remains unknown).

Knowledge of the function and three-dimensional structure of SopD would improve our understanding of the molecular mechanisms involved in *Salmonella* infection. Such understanding is necessary for the development of new compounds to treat *Salmonella*-induced diseases. Here, full-length and truncated versions of *S. dublin* SopD (Fig. 1) have been expressed in an *Escherichia coli* expression system and purified. Biophysical and sequence analyses have been performed. Crystals have been produced for preliminary X-ray analysis. In addition, since truncated SopD contains 12 methionine residues out of a total of 301 amino acid residues, selenomethionine (SeMet) labelling of truncated SopD has been carried out as this provides an attractive alternative to heavy-atom phasing methods in solving crystal structures.

2. Materials and methods

2.1. SopD constructs

Due to its potential to increase solubility and facilitate purification, we used a GST fusion system for expression of both full-length SopD and SopD lacking 16 N-terminal residues (SopDΔN) (Fig. 1). Specifically, both full-length and truncated *sopD* cDNA from *S. dublin* 2229 were cloned into the pGEX-2T expression vector (Amersham Biosciences). The pGEX-2T vector expresses the protein of interest as a fusion to the C-terminus of GST under the control of the *tac* promoter, with a thrombin cleavage site between GST and protein of interest.

2.2. Expression and purification of SopD and SopDΔN

E. coli BL21 Star™ (DE3) cells (Invitrogen) transformed with pGEX-2T plasmids expressing either GST-SopD or GST-SopDΔN were grown in Luria–Bertani (LB) medium with 100 μg ml^{-1} ampicillin at 37 °C to an optical density at 600 nm ($OD_{600\ nm}$) of 0.7–1.0 prior to induction with 0.5 mM isopropyl-β-D-thiogalactoside (IPTG) and incubation for a further 4 h at 30 °C. Bacterial cells were harvested and resuspended in MTPBS lysis buffer (150 mM NaCl, 16 mM Na_2HPO_4, 4 mM NaH_2PO_4, 0.1% Triton™ X-100 and 0.2% β-mercaptoethanol) and lysed by sonication. Cell debris was removed by centrifugation (20,000 × *g*, 20 min, 4 °C) to yield a supernatant containing the soluble GST fusion protein. To capture the fusion protein, the protein lysate was added to glutathione sepharose 4B resin (Amersham Biosciences) and incubated for 1 h at 4 °C on a roller mixer. The resin was recovered by centrifugation (100 × *g*, 5 min, 4 °C) and washed with ice-cold thrombin cleavage buffer (50 mM Tris–HCl (pH 8.0), 150 mM NaCl, 2 mM dithiothreitol (DTT), 5 mM $MgCl_2$, 2.5 mM $CaCl_2$).

SopD or SopDΔN was separated from GST by thrombin digestion: the washed and equilibrated resin was resuspended in thrombin cleavage buffer and 15 U of thrombin (Calbiochem) was added. The resin/buffer slurry was then incubated at 4 °C on a roller mixer for 12 h. Thrombin was removed by incubation with *p*-aminobenzamidine agarose (Sigma). Further purification was achieved by anion exchange (using an ÄKTA Purifier system, Amersham Biosciences) on a Mono Q HR 5/5 column (Amersham Biosciences) equilibrated in 20 mM Tris–HCl, pH 8.0 (the SopD or SopDΔN sample was diluted two-fold using 20 mM Tris–HCl, pH 8.0 in order to reduce the salt concentration prior to anion exchange). Proteins were eluted with a 0–1 M NaCl gradient at a flow rate of 1 ml min^{-1}. Pure SopD or SopDΔN fractions from several anion exchange runs were pooled and concentrated to 10 mg ml^{-1} by ultrafiltration and exchanged into 10 mM Tris–HCl (pH 7.5), 5 mM DTT. Between 20 and 30 mg of pure SopD or SopDΔN was obtained from 1 l of cell culture. The identity of the purified proteins was confirmed using mass spectrometry. SopD and SopDΔN protein concentrations were measured using the theoretically calculated extinction coefficient and by measuring A_{280}.

2.3. Preparation of selenomethionine-labelled SopDΔN

The expression protocol described above was slightly altered to produce SeMet-labelled SopDΔN. SeMet-labelled SopDΔN was expressed and purified using the GST fusion system described in Section 2.1. Because the *E. coli* strain used is not a methionine autotroph,

SeMet labelling was performed by blocking the methionine biosynthesis pathway with high concentrations of exogenous amino acids known to inhibit methionine biosynthesis, following a published protocol with minor modifications [13]. *E. coli* strain BL21 (DE3) transformed with the GST-SopDΔN plasmid was grown overnight at 37 °C in LB medium containing 100 μg ml^{-1} ampicillin. The cells were harvested and resuspended in 1 l M9 minimal medium supplemented with 1 g l^{-1} NH_4Cl, 40 ml 20% (w/v) glucose, 4 ml 0.1% (w/v) D-biotin, 4 ml 0.1% (w/v) thiamine–HCl, 1 ml 1 M $MgSO_4$ and 0.3 ml 1 M $CaCl_2$. The cells were grown at 37 °C until $OD_{600\ nm}$ reached 0.7–0.8 at which point 50 mg l^{-1} L-selenomethionine (Sigma), 100 mg l^{-1} L-lysine-hydrochloride (Fisher), L-threonine and L-phenylalanine, and 50 mg l^{-1} L-leucine, L-isoleucine and L-valine (Melford) were added. The cells were then incubated at 37 °C for 15 min before induction with 0.1 mM IPTG and incubation at 25 °C for a further 6 h. The cells were harvested and the SeMet-SopDΔN was purified using the same methods as for SopDΔN except that 10 mM DTT was added throughout to prevent SeMet oxidation.

2.4. Limited proteolysis

A one-dimensional ^1H NMR spectrum was recorded at 25 °C of 0.2 mM SopDΔN in 10 mM perdeuterated Tris, 5 mM perdeuterated DTT in D_2O, pH 7.5. Proteolysis was then performed by adding 1 μl of 10 mg ml^{-1} trypsin to this NMR sample. Since there was no change in the NMR spectrum at 25 °C after adding trypsin, the sample temperature was subsequently raised to 30 °C (again, there was no change in the NMR spectrum) and then 37 °C. At intervals following trypsin addition, a one-dimensional ^1H NMR spectrum was recorded and a sample taken for sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) in order to monitor changes in the protein due to trypsin cleavage. When the reaction was complete, as assessed by one-dimensional ^1H NMR, the trypsin-resistant core of SopDΔN was subjected to liquid chromatography-electrospray mass spectrometry to identify the site(s) of trypsin cleavage. Before it was used for crystallization trials, the trypsin-digested SopDΔN NMR sample was purified by ion exchange as described in Section 2.2 for SopD and SopDΔN.

2.5. Circular dichroism

Circular dichroism (CD) was performed on a Jasco J-710 spectrometer using a 0.1 cm cell at 25 °C. Temperature was controlled using a water-jacketed cell holder. The concentration of the SopDΔN sample was 2 mg ml^{-1} in 20 mM Tris, pH 8.0 and 5 mM DTT. The system was purged with 22 l min^{-1} nitrogen during data acquisition. Spectra were collected over a wavelength range of 185–260 nm at 0.1 nm resolution and a band width of 2.0 nm. Final spectra were the sum of 10 scans accumulated at a speed of 200 nm min^{-1} with a response time of 0.5 s. Spectra were corrected for contributions from buffer by subtracting the buffer spectra. CD spectra were analysed using the K2D software package [14].

2.6. Analytical size exclusion chromatography

Analytical size exclusion chromatography was carried out at room temperature using an ÄKTA Purifier chromatography system (Amersham Biosciences) on a Superdex™ 75 HR 16/60 column (Amersham Biosciences) equilibrated with 50 mM Tris–HCl (pH 7.5), 100 mM KCl (these are the recommended conditions for the molecular weight standards). SopDΔN was loaded at a concentration of 2 mg ml^{-1}. Protein was eluted over 1.5 column volumes at a flow rate of 0.7 ml min^{-1}. The column was calibrated with BSA (66 kDa), carbonic anhydrase (29 kDa), cytochrome *c* (12.5 kDa) and aprotinin (6 kDa).

2.7. Crystallization and data collection

SopD, SopDΔN and trypsin-digested SopDΔN (tSopDΔN) crystals were grown using vapour diffusion techniques. Initial crystallization trials were performed with sparse matrix kits (Clear strategy I and II, Molecular Dimensions Ltd.). The hanging drop was prepared by mixing 2 μl of 10 mg ml^{-1} SopD, SopDΔN or tSopDΔN with 2 μl of reservoir solution and equilibrated against 1 ml reservoir solution at 4 °C.

X-ray diffraction data from crystals mounted in a quartz capillary tube were collected at room temperature with a MAR Research image-plate detector using Cu Kα radiation from a Rigaku rotating-anode generator. To achieve acceptable statistics, all crystals were exposed for 30 min. Diffraction data were processed using DENZO and SCALEPACK [15].

2.8. Sequence analysis

Secondary structure predictions were conducted using PHDsec [16]. Putative coiled coil domains were identified using the COILS program [17]. A 28-residue window size was used with weighting in favour of hydrophobic residues at the a and d positions of the heptad repeat to avoid false positives (MTIDK matrix). Further analysis was carried out using the MULTICOIL program [18] to determine whether the coiled coil domains were likely to assume dimeric or trimeric structures.

3. Results and discussion

3.1. Expression and purification of SopD and SopDΔN

Full-length SopD derived from *S. dublin* was cloned into the expression plasmid pGEX-2T. Many of the type III

effectors that have been crystallized have had some or all of the N-terminal, non-catalytic portion removed [19–21]. This region is generally unstructured and as such may hinder crystallization. With this in mind, we also generated SopDΔN, with a 16 amino acid N-terminal truncation (Fig. 1), and cloned this into pGEX-2T. These constructs express SopD (approximately 36 kDa) or SopDΔN (approximately 34 kDa) fused to the C-terminus of GST (26 kDa) with a thrombin cleavage site between GST and SopD/SopDΔN, under the transcriptional control of the *tac* promoter. Induction of GST-SopD in *E. coli* BL21 Star™ (DE3) at 30 °C resulted in significant overexpression of a protein that migrates as a 55 kDa band (based on its sequence, GST-SopD is approximately 62 kDa) on Coomassie-stained SDS-PAGE gels and which was absent from uninduced cells (Fig. 2, lanes 1 and 2). When the induced cells were harvested, lysed, and centrifuged to separate soluble and insoluble fractions, GST-SopD was detected mostly (approximately 65%) in the soluble fraction (Fig. 2, lanes 4 and 5).

GST-SopD was initially captured by affinity chromatography using glutathione sepharose 4B resin (Fig. 2, lane 8). SopD was then released from the GST moiety by thrombin cleavage. GST-SopD thrombin cleavage approached completion using an overnight incubation at 4 °C with no secondary degradation products visible by SDS-PAGE, yielding protein with relatively little contamination, as judged by SDS-PAGE (Fig. 2, lane 9). Since protein with purity greater than 95% is required for structural analyses, SopD was further purified using anion exchange chromatography on a Mono Q column. Using a linear 0–1 M NaCl gradient at pH 8.0, the protein eluted at 0.2–0.3 M NaCl. The major elution fraction contained a single 36 kDa band with the other major contaminants, GST and thrombin, eluting off the column separately. Thus ion exchange yielded SopD of greater than 95% homogeneity, as judged by SDS-PAGE analysis (Fig. 2, lane 11). Between 20 and 30 mg of pure SopD was obtained from 1 l of cell culture.

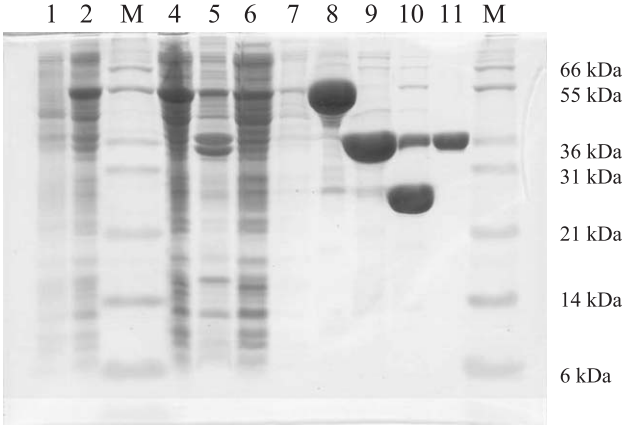

Fig. 2. 15% SDS-PAGE of SopD expression and purification stained with Coomassie brilliant blue. (1) Uninduced *E. coli* BL21 Star™ (DE3) cells; (2) induced *E. coli* BL21 Star™ (DE3) cells expressing GST-SopD; (4 and 5) soluble and insoluble protein fractions after cell lysis; (6) supernatant post-binding to glutathione sepharose 4B resin; (7) glutathione sepharose 4B resin wash; (8) GST-SopD bound to glutathione sepharose 4B resin; (9) supernatant post-thrombin cleavage; (10) glutathione sepharose 4B resin post-thrombin cleavage; (11) pure SopD after anion exchange chromatography; (M) molecular weight markers.

SopDΔN was expressed and purified as described above for SopD. The yield of pure SopDΔN was typically higher than that of SopD, most likely due to the improved solubility (approximately 75%) in *E. coli* of the GST-SopDΔN construct. This, together with the expectation that SopDΔN might crystallize more readily than SopD, meant that most experiments were carried out on SopDΔN rather than SopD.

3.2. Expression and purification of selenomethionine-labelled SopDΔN

Both SopD constructs crystallized readily, but SopDΔN produced better quality crystals and was chosen to be SeMet-labelled for structural determination. The expression of SopDΔN had to be slightly altered to produce SeMet-labelled SopDΔN. Induction at 30 °C with IPTG concentrations higher than 0.35 mM resulted in packaging of most of the fusion protein into insoluble inclusion bodies. This is most probably because the increased hydrophobicity of surface-exposed SeMet residues compared to methionine increases the aggregation tendency.

Cells cultured in M9 minimal medium typically have a doubling time twice that of a richer medium such as LB, and a longer induction time must be used to obtain sufficient quantities of protein. When *E. coli* cells expressing SeMet-SopDΔN were induced for greater than 10 h, however, cell death was observed and very little protein was expressed. This may have been due to metabolism of SeMet into toxic selenium compounds when the cells reached stationary phase, leading to cell death. To circumvent these expression problems, SopDΔN was expressed at 25 °C with 0.1 mM IPTG for 6 h. SeMet-SopDΔN was purified using the same protocol as for the native protein except for the addition of 10 mM DTT throughout the process to prevent SeMet oxidation. About 10 mg of pure SeMet-SopDΔN was obtained from 1 l of cell culture.

3.3. SopDΔN is a monomer in aqueous solution

Analysis by analytical size exclusion chromatography showed that SopDΔN forms a single species consistent with a monomer of 34–35 kDa (data not shown), as calculated from the line for a set of globular protein standards. This compares with a mass of 34,246 Da predicted by the amino acid sequence. Since analytical size exclusion chromatography is sensitive to both molecular mass and shape of the solute, the excellent match between predicted and experimentally determined values suggests that SopDΔN is a monomeric, globular molecule.

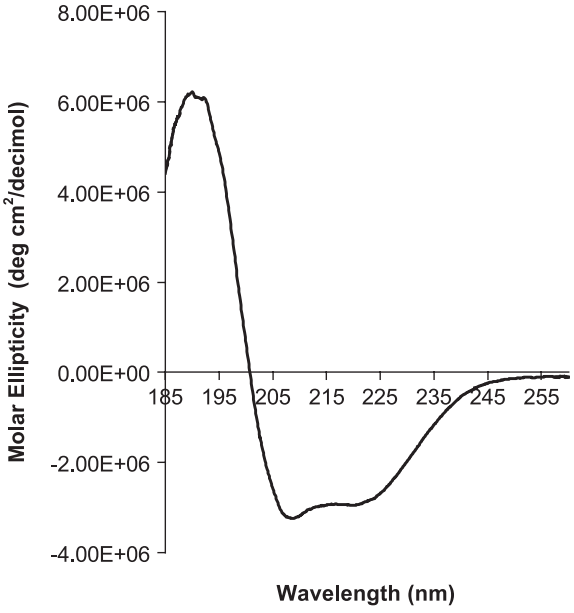

Fig. 3. CD spectrum of SopDΔN showing that the protein has mixed secondary structure composition with a higher percentage of α-helix than β-strand.

3.4. SopD sequence analysis

Analysis of the protein sequence using the COILS server http://www.ch.embnet.org/software/COILS_form.html) highlighted the presence of a putative coiled coil motif involving residues 271–298, suggesting that the protein may self-associate or associate with one or more other proteins. Further analysis using the MULTICOILS server http://multicoil.lcs.mit.edu/cgi-bin/multicoil) produced a very low probability (<0.1%) for dimer or trimer formation. This result, combined with the indication from size exclusion chromatography (see above) that SopD is monomeric, means that any interactions mediated by the coiled coil motif are likely to be heterotypic rather than homotypic, although it should be emphasized that no binding partners of SopD have yet been identified.

The coiled coil is a widespread and versatile motif found in a variety of structural and regulatory proteins [22]. Numerous other type III-secreted proteins possess coiled coil motifs [23]. Coiled coil-containing proteins associated with TTSSs can be placed into three broad, function-based categories [23]: (i) proteins involved in regulation of secretion; (ii) secreted translocators (proteins involved in the formation of structures on the host cell or bacterial cell membrane that are required for passage through the host cell membrane of other effector proteins and possibly translocators themselves); (iii) proteins with enzyme activity. SopD is neither a regulator of secretion nor is it a secreted translocator, and any capacity of SopD to function as an enzyme remains to be identified.

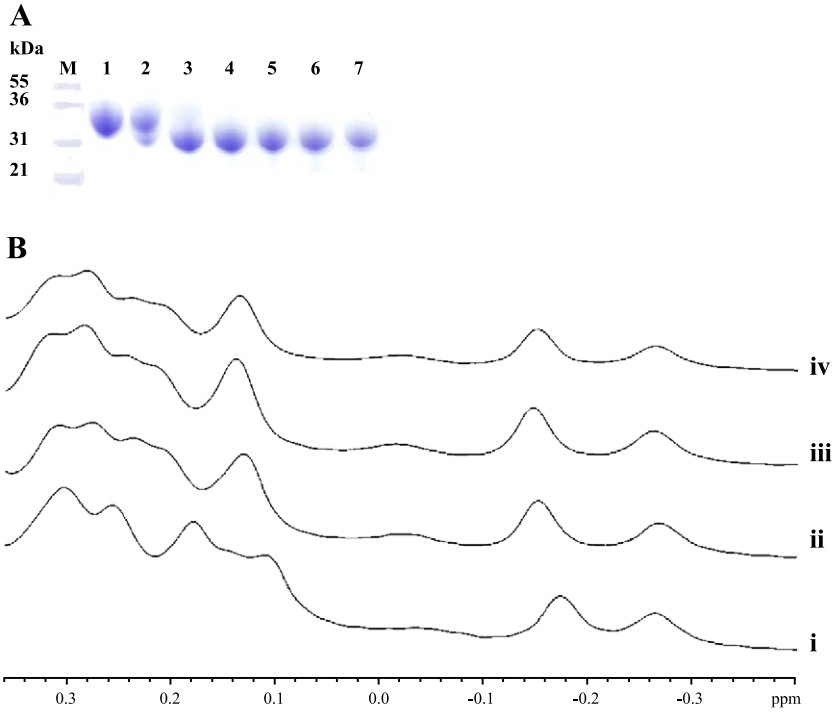

Fig. 4. Limited trypsin digestion of SopDΔN. (A) Trypsin digestion time course monitored by 15% SDS-PAGE. (M) Molecular weight markers. Lanes 1–7: 0, 40, 60, 80, 100, 150, 180 min after raising the temperature of the NMR sample to 37 °C (note that there was no change in the NMR spectrum at 25 or 30 °C after addition of trypsin). (B) Trypsin digestion time course monitored by ^{1}H NMR spectroscopy. Upfield region of the ^{1}H NMR spectrum of SopDΔN (i) 0; (ii) 60; (iii) 100; (iv) 180 min after increasing the sample temperature to 37 °C.

3.5. SopD has a mixed secondary structure composition

The far-UV CD spectrum of SopD exhibits minima at 208 and 222 nm that are characteristic of α-helices (Fig. 3). Calculation of secondary structure based on the CD spectrum using the K2D software package [14] indicates that SopD comprises 38% α-helix and 26% β-strand. This is in reasonable agreement with a theoretical prediction that SopD is composed of 46% α-helix and 14% β-structure [16]. Both experiment (CD spectroscopy) and theory (the PredictProtein server at http://www.embl-heidelberg.de/predictprotein/predictprotein.html) therefore indicate that SopD has a mixed secondary structure composition with α-helix in greater abundance than β-strand.

3.6. Domain delineation by limited proteolysis

Delineation and sub-cloning of the protease-resistant core domain of a protein often results in a polypeptide that is more amenable to crystallization and/or NMR studies. Limited trypsin proteolysis was consequently carried out on an NMR sample of SopDΔN (SopD was not chosen for this experiment because an NMR sample of SopD had previously exhibited extensive precipitation). The time-course of proteolysis was monitored by both SDS-PAGE (Fig. 4A) and ^{1}H NMR spectroscopy (Fig. 4B). SDS-PAGE can report on size changes in a protein due to proteolysis and the ^{1}H NMR spectrum is very sensitive to local changes in a protein due to proteolysis. The end-point of the proteolysis reaction can be determined in real time by this method because there is no further change in the NMR spectrum when the reaction is complete.

There was no change in the NMR spectrum after trypsin addition over a period of 2 h when the temperature was set to 25 °C and then 30 °C, but on increasing the temperature to 37 °C, slight changes were observed in the ^{1}H NMR spectrum (Fig. 4B). A ^{1}H NMR spectrum was recorded and a sample taken for SDS-PAGE at 40, 60, 80, 100, 150 and 180 min after raising the temperature of the sample to 37 °C. The SDS-PAGE band (Fig. 4A) and ^{1}H NMR spectrum (Fig. 4B) remained essentially constant between 60 and 180 min after raising the temperature to 37 °C. Since this indicates that the SopD core domain produced by trypsin digestion (tSopDΔN, Fig. 1) is quite resistant to further proteolysis, it was decided that, rather than sub-clone a new construct corresponding to the protease-resistant core domain, trypsin digestion of SopDΔN would be incorporated into the original SopDΔN purification protocol in order to generate tSopDΔN. This has been achieved with excellent results (data not shown), although the poor initial crystallization trial results for tSopDΔN (Section 3.7 below) meant that tSopDΔN has not been generated routinely.

The trypsin cleavage site was identified by liquid chromatography-electrospray mass spectrometry to be between residues K14 and A15 of SopDΔN, corresponding to K30 and A31 of full-length SopD (Fig. 1). These studies thus revealed that a domain spanning residues 31–317 is stable for at least 2 h in the presence of trypsin at 37 °C (Fig. 4). This suggests that SopD comprises a single compact domain with some of the putative N-terminal type III secretion signal exposed. According to recent evidence for a universal mode of interaction between TTSS chaperones and effectors that involves independent chaperone binding and effector domains [24], the single-domain architecture of SopD suggests that this protein does not have a cognate bacterial chaperone. The same authors, furthermore, proposed a model for temporal control of secretion [24] according to

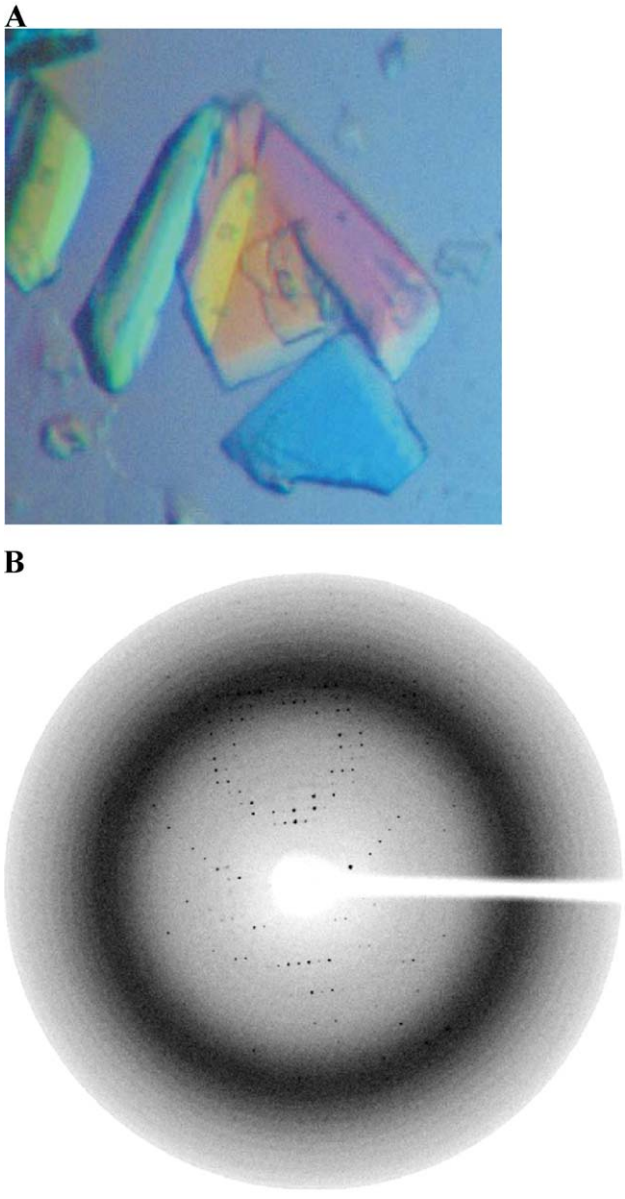

Fig. 5. Crystallization and X-ray diffraction of *S. dublin* SopDΔN. (A) Crystal of SopDΔN grown at 19 °C using the hanging drop method. The dimensions of the largest crystals were 0.1 × 0.1 × 0.2 mm. (B) Diffraction pattern of SopDΔN crystals. The resolution at the edge of the plate is 2.98 Å.

which SopD would be secreted in the second group of effectors following effectors with cognate chaperones.

3.7. Crystallization and data collection

Initial crystallization screens using purified SopD and SopDΔN (10 mg ml^{-1}) yielded needle-like crystals in a variety of conditions with KBr and PEG as precipitants at pH 6.5. These crystals grew within hours but were too small to mount and their diffraction quality could not be assessed. An initial crystallization screen using tSopDΔN was not successful, with several conditions resulting in protein precipitation. SopDΔN crystallized more readily under a greater variety of conditions than SopD and was chosen to optimise crystallization conditions. Attempts to increase the crystal size by altering the temperature, protein concentration and precipitant had no discernible effect. Diffraction quality crystals were obtained with the addition of phenol (10–20 mM) and dioxane (5–10%) with 15% PEG 6K and 0.15 M KBr as precipitants buffered with 100 mM PIPES, pH 6.5. Under these conditions, crystals grew to their maximum size ($0.1 \times 0.1 \times 0.2$ mm; Fig. 5A) in 3 days and diffracted X-rays to 3.0 Å on our 'in-house' X-ray source (Fig. 5B). Crystals belonged to space group $P2_1$, with unit-cell parameters $a = 67.9$ Å, $b = 81.8$ Å, $c = 77.1$ Å, $\alpha = 90.0°$, $\beta = 107.0°$ and $\gamma = 90.0°$, and contained two SopDΔN molecules in the asymmetric unit with a solvent content of 60%. Attempts at numerous flash-freezing protocols using PEG 400, glycerol, ethylene glycol and methyl pentanediol as a cryoprotectant severely reduced the crystals' ability to diffract; higher quality data could only be collected at room temperature with the crystal mounted in a quartz capillary tube.

4. Conclusions

Since SopD does not exhibit significant sequence homology to any other protein in the databank except SopD2 [12], it has proven difficult to establish the precise biochemical function(s) of SopD in *Salmonella* pathogenesis. Structural information may assist in defining SopD function. To that end, we have undertaken expression, purification, biophysical characterization and initial crystallization studies on *S. dublin* SopD and SopDΔN. We have established that SopDΔN is monomeric in solution, of mixed secondary structure composition, and that it comprises a short, protease-susceptible N-terminal region and a much larger protease-resistant core domain. The single-domain architecture suggests that SopD is unlikely to have a cognate chaperone [24]. High SopD sequence identity between different *Salmonella* serovars means that these findings are likely to apply to SopD proteins from most, if not all, *Salmonella* serovars. Finally, we have made good progress in defining conditions under which crystals of sufficient quality can be produced to determine the three-dimensional structure of SopD.

Acknowledgements

This work was supported at Bath by The Wellcome Trust (grant 060998 to S.B.) and the Royal Society (grant 22604 to J. vd E.) and at IAH by BBSRC. The Wellcome Trust is acknowledged for purchase of the 600 MHz NMR spectrometer (grant 051902) used in this study. C.W. is supported by a PhD studentship from EPSRC.

References

[1] B.D. Jones, S. Falkow, Salmonellosis: host immune responses and bacterial virulence determinants, Annu. Rev. Immunol. 14 (1996) 533–561.
[2] M.E. Ohl, S.I. Miller, Salmonella: a model for bacterial pathogenesis, Annu. Rev. Med. 52 (2001) 259–274.
[3] G.R. Cornelis, F. Van Gijsegem, Assembly and function of type III secretory systems, Annu. Rev. Microbiol. 54 (2000) 735–774.
[4] C.J. Hueck, Type III protein secretion systems in bacterial pathogens of animals and plants, Microbiol. Mol. Biol. Rev. 62 (1998) 379–433.
[5] T. Kubori, Y. Matsushima, D. Nakamura, J. Uralil, M. Lara-Tejero, A. Sukhan, J.E. Galán, S. Aizawa, Supramolecular structure of the *Salmonella typhimurium* type III protein secretion system, Science 280 (1998) 602–605.
[6] T. Kubori, A. Sukhan, S.I. Aizawa, J.E. Galán, Molecular characterization and assembly of the needle complex of the *Salmonella typhimurium* type III protein secretion system, Proc. Natl. Acad. Sci. U. S. A. 97 (2000) 10225–10230.
[7] A. Blocker, N. Jouihri, E. Larquet, P. Gounon, F. Ebel, C. Parsot, P. Sansonetti, A. Allaoui, Structure and composition of the *Shigella flexneri* 'needle complex', a part of its type III secreton, Mol. Microbiol. 39 (2001) 652–663.
[8] F.S. Cordes, K. Komoriya, E. Larquet, S.X. Yang, E.H. Egelman, A. Blocker, S.M. Lea, Helical structure of the needle of the type III secretion system of *Shigella flexneri*, J. Biol. Chem. 278 (2003) 17103–17107.
[9] M.A. Jones, M.W. Wood, P.B. Mullan, P.R. Watson, T.S. Wallis, E.E. Galyov, Secreted effector proteins of *Salmonella dublin* act in concert to induce enteritis, Infect. Immun. 66 (1998) 5799–5804.
[10] S.P. Zhang, R.L. Santos, R.M. Tsolis, S. Stender, W.-D. Hardt, A.J. Bäumler, L.G. Adams, The *Salmonella enterica* serotype *typhimurium* effector proteins SipA, SopA, SopB, SopD, and SopE2 act in concert to induce diarrhea in calves, Infect. Immun. 70 (2002) 3843–3855.
[11] M. McClelland, K.E. Sanderson, J. Spieth, S.W. Clifton, P. Latreille, L. Courtney, S. Porwollik, J. Ali, M. Dante, F.Y. Du, S.F. Hou, D. Layman, S. Leonard, C. Nguyen, K. Scott, A. Holmes, N. Grewal, E. Mulvaney, E. Ryan, H. Sun, L. Florea, W. Miller, T. Stoneking, M. Waterston, R. Waterston, R.K. Wilson, Complete genome sequence of *Salmonella enterica* serovar *typhimurium* LT2, Nature 413 (2001) 852–856.
[12] J.H. Brumell, S. Kujat-Choy, N.F. Brown, B.A. Vallance, L.A. Knodler, B.B. Finlay, SopD2 is a novel type III secreted effector of *Salmonella typhimurium* that targets late endocytic compartments upon delivery into host cells, Traffic 4 (2003) 36–48.
[13] G.D. Vanduyne, R.F. Standaert, P.A. Karplus, S.L. Schreiber, J. Clardy, Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin, J. Mol. Biol. 229 (1993) 105–124.
[14] M.A. Andrade, P. Chacon, J.J. Merelo, F. Moran, Evaluation of secondary structure of proteins from UV circular dichroism spectra using an unsupervised learning neural network, Protein Eng. 6 (1993) 383–390.
[15] Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode, Methods Enzymol. 276 (1997) 307–326.

[16] B. Rost, PHD: predicting one-dimensional protein structure by profile-based neural networks, Methods Enzymol. 266 (1996) 525–539.
[17] A. Lupas, Predicting coiled-coil regions in proteins, Curr. Opin. Struct. Biol. 7 (1997) 388–393.
[18] E. Wolf, P.S. Kim, B. Berger, MultiCoil: a program for predicting two- and three-stranded coiled coils, Protein Sci. 6 (1997) 1179–1189.
[19] G. Buchwald, A. Friebel, J.E. Galán, W.-D. Hardt, A. Wittinghofer, K. Scheffzek, Structural basis for the reversible activation of a Rho protein by the bacterial toxin SopE, EMBO J. 21 (2002) 3286–3295.
[20] A.G. Evdokimov, J.E. Tropea, K.M. Routzahn, D.S. Waugh, Crystal structure of the *Yersinia pestis* GTPase activator YopE, Protein Sci. 11 (2002) 401–408.
[21] C.E. Stebbins, J.E. Galán, Maintenance of an unfolded polypeptide by a cognate chaperone in bacterial type III secretion, Nature 414 (2001) 77–81.
[22] A. Lupas, Coiled coils: new structures and new functions, Trends Biochem. Sci. 21 (1996) 375–382.
[23] R.M. Delahay, G. Frankel, Coiled-coil proteins associated with type III secretion systems: a versatile domain revisited, Mol. Microbiol. 45 (2002) 905–916.
[24] S.C. Birtalan, R.M. Phillips, P. Ghosh, Three-dimensional secretion signals in chaperone-effector complexes of bacterial pathogens, Mol. Cell 9 (2002) 971–980.